Candidates for within-vehicle auditory displays

Denis McKeown
Institute of Psychological Sciences
University of Leeds
Leeds LS2 9JT
United Kingdom
(*d.mckeown@leeds.ac.uk*)

ABSTRACT

Speech, auditory icons (sounds imitating real world events), environmental sounds (naturally occurring sounds), and abstract warnings are all candidates for user interfaces. Such auditory displays and warnings for within-vehicle use must satisfy certain criteria such as being appropriately urgent, acceptable to users and commanding accurate and appropriately fast response times. Here such criteria are investigated and compared for these different forms of auditory display as an interface for a broad range of driving scenarios. In a computer task of identifying learned mappings of sound to scenario, speech and auditory icons produced both faster response times and greatest accuracy. Abstract sounds produced the slowest response times and least accuracy. Environmental sounds showed an intermediate pattern of performance for accuracy but the response times were similar to the abstract sounds. Urgency and pleasantness judgments showed an interesting contrast. Speech utterances were similarly and consistently rated as pleasant, but also of intermediate urgency (that is, speech sounds did not differ according to situational urgency). On the other hand the three other sound types mapped successfully onto their specified situational urgency levels, and showed a consistent relationship: sounds mapped to highly urgent scenarios were also judged less pleasant.

1. INTRODUCTION

Within the Institute of Psychological Sciences at Leeds University, research currently addresses the *appropriate* uses of auditory displays and warnings, as well as their *pleasingness*, for future application in vehicles [1]. A number of developments of Intelligent Transportation Systems (ITS) have been introduced, or are being developed, in order to reduce accidents (as well as contribute to a more efficient transportation system). These use advanced technologies to monitor the traffic environment, as well as calculate and control appropriate driver responses. Auditory displays promise the obvious benefit of not competing directly with the primarily visuo-spatial information processing activities of the driving task itself. Not surprisingly, auditory warnings may attract a person's attention faster than visual displays are able to, and thus may result in faster response times than do visual stimuli [2]. Of course auditory displays cannot replace visual displays within the vehicle (apart from anything else, people who are deaf are drivers too), but their appropriate use may offer significant advantages. Deatherage [3] suggests that auditory displays will be most effective when the message is short, simple, will not be referred to later, deals with events that are not spatial in nature, requires immediate action, and the visual system is overburdened or the receiving location is too bright or too dark. Conversely visual displays are most likely to be most effective when the message is complex, long, will be referred to later, deals with locations in space, does not require immediate action, and the auditory system is overburdened or the receiving location is too noisy.

As well as speech messages and abstract noise and tone warnings, auditory displays can use 'earcons,' which are short musical sequences with meanings that may be learned by the user, or 'auditory icons,' which are characteristic sounds signifying real world objects or events. For example a rumble sound may indicate straying from the highway, or a horn sound may indicate the required action to brake. Auditory icons are frequently environmental or everyday sounds, since these already have strong metaphorical mappings to the events they signify. As with visual icons, auditory icons have originated from the idea that by making them representational users will have a more direct feeling of engagement with the system, with such sounds providing intuitive links to the objects or events that they represent. Crucially, learning, identifiability, and recall of auditory icons may benefit from such links when compared to abstract displays or warnings. Indeed, recently a taxonomy of auditory icons has been proposed [4], with learning benefiting particularly from what the authors term 'direct relations' between sound and referent (typically the sound is that made by the single signified event in the real world). One encouraging result is reported by Graham [5], who examined the effectiveness of auditory icons (a car horn sound, a tyre screeching sound) in a collision warning scenario within a driving simulator. Comparison was made with tonal and speech warnings. Participants were required to decide whether to brake or not depending on the warning and the scene presented to them. The auditory icons produced significantly faster response times than either the tone or speech warnings. Although Graham does allow that more complex abstract sounds, such as the synthetic musical sequences known as earcons, might produce comparable reaction times to the auditory icons, he noted that such sounds may not produce such an ease of interpretation or the same reduced confusability.

Auditory icons may be easy to understand when people are first presented with them. Begault [6] found that auditory icons can be classified and categorised by users better than non-representational sounds. Stanton and Edworthy [7] have advocated a user-centred design approach based upon the notion of affordances [8] where auditory displays signify actions for users. These authors emphasise that auditory affordances may depend critically on the context in which the auditory icons occur, as well as the prior experience of users. For example, they found that new staff within an intensive treatment unit more readily acquired the appropriate response to representational sounds (such as bubbles representing a syringe pump) than old staff already familiar with existing conventional abstract warning sounds in the unit.

More often than not explorations of representational auditory displays have been in the area of human-computer interaction. For example, sounds representing computer command functions were compared in one study; here, although speech was most easily identified and was preferred by participants, auditory icons were associated with the correct command function more than earcons [9]. Speech, auditory icons and earcons were also compared in an abstract association task [10]; here speech produced the fastest association times and fewest errors.

1.1. Perceived urgency of auditory displays

Urgency is a critical consideration in the design of auditory warnings and displays. Sounds generally contain inherent urgency levels (either based upon acoustic features, learned associations, or their interaction), so that listeners can evaluate the urgency of a sound even if the meaning is unclear or not known. If not appropriately scaled for level of urgency an operator could respond to a routine non-critical alarm because it sounds urgent, at the potential expense of another more urgent reaction that is required elsewhere, or conversely, an operator might not respond to a critical alarm because it sounds less urgent and seems unimportant [11]. It has been found that the usual simple behavioural measure of response time decreases as perceived urgency increases [12; 13].

Importantly, urgency and appropriate response may be related. In the driving simulator study by Graham [5], using a car horn sound and the sound of screeching tyres, it was found that there was a higher likelihood of participants responding incorrectly to the auditory icons compared to the other warning sounds. Graham suggested that this higher rate of false alarms with auditory icons was due to the high perceived urgency of the auditory icons, resulting in participants not taking enough time to assess the road situation for themselves. Of course any conclusion is limited to the couple of sounds used in the Graham study.

Another issue is the difficulty of manipulating levels of urgency for auditory icons, in contrast to abstract sounds for which researchers have been able to define parameters of the sounds that should be manipulated in order to alter urgency levels [14; 15; 16]. But the identification of acoustic properties governing the level of perceived urgency is unlikely to be exhaustive; in any case there may be no substitute for running a series of tests of urgency with any new set of candidate sounds. Obviously, changing icon sounds acoustically (or just speeding up their delivery) may well affect their meaningfulness. More useful may be simply to choose the sounds according to their intrinsic or learned urgency, the approach adopted in the present study. But it is recognised that context of delivery may be a critical factor [17].

1.2. User acceptance of auditory displays

It can hardly be overstated that auditory displays within vehicles should be acceptable to the user. Sounds that are considered to be unpleasant or annoying would probably be disabled by the driver. And alarms may produce unwelcome startle responses in users [18]. Of course, in some situations unpleasantness may be desirable. Lazarus and Höge [19] noted that sirens and horns, two sounds that could potentially be used as within-vehicle auditory warnings, were incompatible with pleasant situations. They noted that "the greater the difference of compatibility between the danger signals and pleasant situations the better" (p.46).

The majority of research carried out into unpleasantness or annoyance and sound has focused on the physical characteristics of the sound which are considered to be unpleasant. Laird and Coye [20] observed a close relationship between annoyance and loudness, with high-pitched sounds being intrinsically more annoying than medium- or low-pitched sounds. Spieth [21] argued that generally, annoyance of sound does not vary with frequency when differences in loudness are accounted for, with the exception of sounds that are above 6 kHz, which are moderately more annoying then sounds of lower frequency. The cognitive context in which sounds are heard may affect annoyance [22].

1.3. Sound-referent mappings

The inherent meaningfulness of auditory icons may be particularly important [4]. For example, an emergency situation is likely to occur infrequently, so that the driver will not often be exposed to an alarm. Thus, the opportunity to learn its referent may be limited in everyday use; therefore a sound that carries inherent meaningfulness in this situation would seem ideal. But as noted above, a possible disadvantage of auditory icons is that acoustic manipulations of such sounds to achieve a certain level of situational urgency may make them unrecognisable. This points to the importance of selecting auditory icons that have appropriate urgency, and also have appropriate meanings, for selected within-vehicle situations. In the present study a range of what may be termed environmental sounds were selected (for example, rainfall, baby crying) and their perceived urgency and perceived unpleasantness were measured; then an assessment was made of how these two perceptual dimensions related to one another.

Of course, the goal of mapping a sound to its referent may be advanced by manipulating appropriate acoustic characteristics [15]. This might be done whilst preserving the inherent meaning of the sound. Intuitively a sound that is loud, fast-repeating and with dominating high-frequency content or inharmonic structure may command a more urgent response, or more urgent judgment in the listener; and indeed empirically this appears to be the case [16]. Recently, Guillaume and colleagues [17, Experiment 1] asked one group of listeners to rate 13 sound sequences [very similar to those used in 15] for their acoustic dissimilarity, and another group to rate them for their urgency; a multi-dimensional scaling analysis showed remarkable concordance between these dimensions. These authors also identified a feature that may take precedence, namely how listeners rapidly encode sounds they are familiar with, such as the rapidly repeating traditional emergency vehicle alarm; they argued that a more exhaustive acoustical appreciation by the listener may actually impair response times.

1.4. Design considerations within vehicles

Auditory displays may be designed to serve a variety of functions within vehicles; for example, they may: alert the driver to a possible hazard; provide an alarm about a hazard that is occurring and requiring immediate action; and warn the driver that a hazard may occur should certain circumstances prevail (these distinctions are discussed in 23). Mappings between differing forms of auditory display (such as speech, auditory icons and abstract alarms) and different hazard scenarios that might be signaled by a within-vehicle collision avoidance system (CAS), require study. Intuitively, abstract alarms may be the most appropriate choice for highest-priority signals (imminent collision with a vehicle in the blind spot, for example following the initiation of a lane change maneuver), since they cannot be easily confused with background enviromental or speech sounds. On the other hand frequenctly employed alerting (for example driving with door open) or advisory functions (tyre pressure is low, for example) might be better served by speech and auditory icons [4]. But it might be argued that speech, too, is perfectly suitable as a high-priority warning [24]. It is one purpose of the present study to help make the move from such intuitive mappings to mappings based on user behaviour.

Environmental sounds – whether running water or the bell of the old-style telephone – would seem to possess strong familiarity and so offer the opportunity for rapid processing by listeners (Gaver and colleagues make a similar case in their discussion paper on auditory icons[1]). Recently, a short study here at Leeds investigated a range of sounds varying from the sound of a baby gurgling to bird song, as candidate within-vehicle auditory displays. Listeners rated each of twenty environmental sounds on 7-point perceived urgency and 7-point unpleasantness scales. Highly significant positive correlations were observed between these two dimensions: that is, sounds that had high urgency were also judged unpleasant. The inference is that, although environmental sounds may perform well when conveying the meaning of the events they signify, user acceptance may be low, at least for those sounds serving as highly urgent alarms (incidentally, Weise and Lee, in press[2], provide a useful recent discussion of this issue).

The present paper investigates speech, a range of environmental sounds (such as footsteps and seashore lapping), abstract sounds (a bell alarm for example), as well as auditory icons (which differed from the environmental sounds in being matched to particular outcomes), as candidates for auditory displays within the vehicle. But here the earlier work is extended to include specific driving scenarios in a response time task, and also employs a free-modulus magnitude estimation procedure, described below, for judged pleasantness and urgency.

2. METHODS

2.1. Participants

Ten participants (4 males, 6 females) took part in an evaluation of situational urgency. Their average age was 40, and average years of driving was 17.1. Forty different participants took part in the main experimental study, with ten each being assigned randomly to one of four groups (speech, auditory icon, environmental and abstract sounds). Average age was 24.05 years and 33 females and 7 males took part. Participants either had normal or corrected to normal vision. All participants reported no known hearing problems with the exception of one participant who reported mild tinnitus.

2.2. Assessing situational urgency

This study was conducted to collect experienced drivers' judgements of a range of driving scenarios. The drivers were required to rate 20 driving scenarios on a scale of 1 to 3; where 1 represented a very urgent 'warning' scenario with milliseconds or seconds to respond; 2 a medium urgency 'alerting' scenario with minutes to respond; and 3 a relaxed 'advisory' scenario with 10 or more minutes to respond. The 20 scenarios were separated into advisory, alerting and warning as a result of the experienced drivers' judgements. Of course, the assignment into just three categories could be viewed as quite arbitrary (auditory displays could be designed to serve for a broad range of situations); what was important was that the scenarios differed progressively and consistently in their judged urgency (that is, our assignments were based upon participants' evaluations of situational urgency). Three scenarios from each of these urgency levels were then chosen to be used in the experimental task.

2.3. Stimuli

A softly spoken English female voice with clear articulation was used for the speech condition. There is some slight evidence that a female voice is more

[1] A draft manuscript available at http://www.billbuxton.com/Audio.TOC.html

[2] A copy may be requested from john-d-lee@uiowa.edu

acceptable to users [24] and therefore more likely to be used in the auditory interface. Each speech message consisted of either 3 or 4 words that described each of the driving scenarios. Each message began with a different word to the other messages in order that no two messages would start with the same word, which would have potentially increased response times for those particular messages. The messages lasted for 3 – 4 seconds and were spoken at a similar pace. But a critical consideration was that the particular scenario was indicated where possible in the first word of each utterance, so that in the experimental response-time tests listeners would not have to listen to the end of each utterance before responding (and indeed the response times indicate they did not do so). Examples are "headway closing fast" and "oil is low." High fidelity recordings were made using a condenser microphone (MBC model 603) within an Industrial Acoustics Company (IAC) sound-isolating booth, with the signal externally sampled at 44 kHz over a Digigram soundcard (model VX Pocket 440) onto a PC using Cubase sound-recording software.

The 'abstract' category included sounds that were not thought to convey any type of meaning to listeners, such as pulses, tones, bells, beeps, and buzzes. The sounds, separated into relaxed, slightly relaxed and urgent on the basis of urgency judgements in a previous experiment in our laboratory, were matched to the three situational urgency categories by the experimenter.

The category of 'environmental sounds' consisted of real-world sounds that were likely to be familiar to people, but did not have specific meanings within the vehicle interface. Environmental sounds were assigned to this experiment based on 7-point urgency rating judgements. They were considered to be distinctive from one another. As with the abstract sounds' category, these sounds were paired with scenarios on the basis of their perceived urgency, with the constraint that they did not specifically represent any of the driving scenarios they were matched to (or indeed any other scenarios encountered in the vehicle).

Like the environmental sounds, the 'auditory icons' were representative sounds, but in this case chosen to resemble as closely as possible particular scenarios within the vehicle; for instance the sound of a car speeding past was chosen to indicate "exceeding speed limit." These sounds were therefore not paired with each scenario due to a matching of situational with perceived urgency, as with the abstract and environmental sounds.

2.4. Procedure

2.4.1. Computer task

Participants were allocated to one of four conditions (speech, auditory icons, environmental, and abstract). This between-groups allocation for the sound groups factor was thought necessary, since familiarity with one class of auditory displays is known to limit learning about another class [7]. The participants' task was to match each sound presented with one of 9 driving scenarios presented pictorially on a computer monitor. Sounds were played to participants at a comfortable listening level through BBC design LS3/5A monitor loudspeakers in the IAC sound-isolating booth. Sound level measurements within the booth ensured the sound stimuli did not differ in overall level (detailed acoustic analyses, not reported here[3], were also carried out). The nine driving scenarios, selected on the basis of participants' responses in the pilot study, were represented simultaneously on a computer monitor in black and white, each approximately 27 mm by 27 mm, and displayed in a semi-circle at an equal distance (110 mm) from a centre circle. A red centre circle (diameter 15 mm) was present at the beginning of each trial. The action of the participant moving the mouse cursor to the centre circle activated the 9 pictures on screen and started the sound, with participants' response times being recorded from this time. Each trial and response time measurement ended when the participant clicked on one of the 9 pictures using the mouse. The pictures appeared in a random order on the screen on each trial. For each of the 4 sound groups, 108 trials occurred in a session (12 random presentations of each of the 9 sounds, with the restriction that no sound occurred twice in succession).

2.4.2. Free modulus magnitude judgements

Two experimental sessions were conducted with each session being exactly 7 days apart. For the first session participants listened to 9 sounds (in a random order) and rated them for pleasantness. Each sound was rated by assigning high numbers to highly pleasant sounds and low numbers to unpleasant sounds, based on a method of magnitude estimation described by Engen [25], and which may be traced to Stevens [26]. It is a direct scaling approach in which an observer attempts to match numbers to perceptions without imposed restrictions: the observer uses those numbers which they feel are appropriate. It is also a standard (or 'modulus') free method – on each stimulus presentation the observer assigns a number to the stimulus, though the instructions request that the ratios between assigned numbers correspond as closely as possible to perceived ratios for the stimulus attribute being assessed. As well as discussing the rationale and example data for the method, Engen also describes a series of steps in converting pairs of assigned numbers per stimulus into a table of least-squares normalized values.

After scoring all 9 sounds for pleasantness, participants were required to do the same for the urgency ratings; this time high numbers represented high urgency and low numbers represented low urgency. This session concluded with participants completing a second scoring of all 9 sounds for pleasantness and then for urgency. For all of the scorings participants were told to concentrate on sound rather than meaning.

[3] This work is the subject of a paper under review in Human Factors, where these analyses are reported

The experimenter provided a demonstration session of the computer task in which each of the 9 sound-scenario pairings were displayed once and the participants were informed of the correct matching of each pairing. Participants then completed two practice sessions where the 9 sound-scenario pairings were presented once each; they were encouraged to ask the experimenter questions during this practice and would be told of the correct pairings whenever they gave incorrect responses. On completion of the practice session participants were informed that their speed and accuracy would be measured throughout the experiment.

Each trial began when the participants moved the mouse cursor to the centre circle, thus activating the sound, on-screen display and response time measurement. A correct sound-picture matching resulted in the picture flashing once before the beginning of the next trial. The experimental session consisted of a block of 108 trials.

Participants returned for the second experimental session one week after they had taken part in the first. This session began with the computer task (108 trials), although in this case participants were not given a demonstration or a series practice trials before starting the task (so that in this case recall from week 1 was being assessed). Once they had completed the computer task participants scored the sounds using free modulus estimations for pleasantness and urgency, twice each; urgency ratings always following pleasantness ratings.

3. RESULTS

The following results show, separately, the outcomes for computer task performance for accuracy and speed of response, and for the free modulus magnitude estimations of pleasantness and urgency.

3.1. Computer task

The *accuracy* of picture selection in the computer task was highest for the speech stimuli, with the auditory icons showing a comparable high level of performance (Figure 1). Lowest accuracy in the task was observed for the abstract sounds, though both these and the environmental sounds showed notable improvements from the first weekly test to the second weekly test. A 2-way ANOVA for mean performance accuracy with factors of recall (week 1, week 2) and sound group (speech, auditory icons, environmental sounds, abstract sounds) showed significant effects for recall ($F(1, 36) = 15.327$, $p < .01$), and sound group ($F(3, 36) = 48.143$, $p < .01$), and a significant interaction between these factors ($F(3, 36) = 4.559$, $p < .01$). Newman-Keuls tests revealed abstract and environmental differed from each other, and these differed from auditory icon and speech groups (ps < .05). The speech and auditory icon sound groups were not significantly different from each other.

The *response times* for picture selection in the computer task were lowest for speech and auditory icons, and highest for environmental and abstract sounds (Figure 2). Slight improvements in response times overall, possibly due to practice, were evident in week 2.

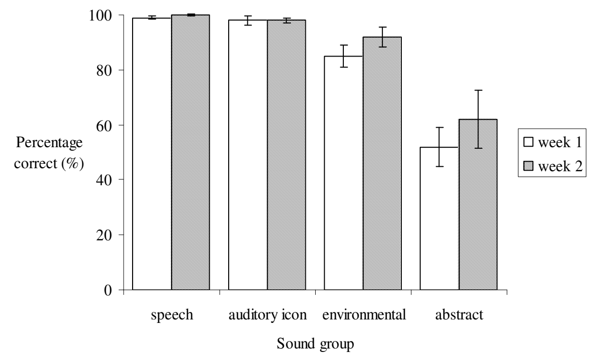

Figure 1. *Mean accuracy in the computer task in weeks 1 and 2.*

A 2-way ANOVA examining mean *speed of responding* with factors of recall (week 1, week 2) and sound group (speech, auditory icons, environmental sounds, abstract sounds) showed a significant effect of recall ($F(1, 36) = 17.364$, $p < .01$) and group ($F(3, 36) = 11.051$, $p < .01$), but no interaction ($p > .1$). Newman-Keuls analyses showed that the environmental and abstract sound groups were significantly different from the speech and auditory icons groups (ps < .05).

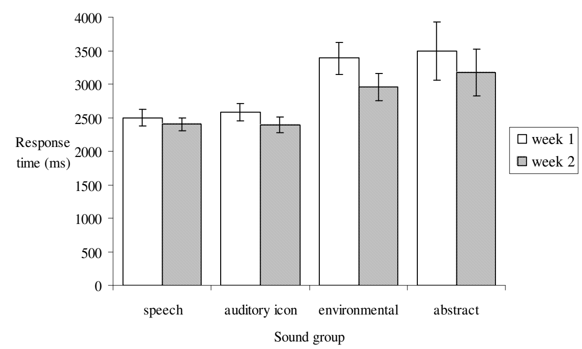

Figure 2. *Mean response times in the computer task in weeks 1 and 2.*

3.2. Free modulus magnitude estimations

As shown in Figure 3 for week 1, auditory icons, environmental and abstract sounds were scored less pleasant as scenarios became more urgent. Speech showed a flat function for pleasantness. Scores for urgency, week 1, increased for auditory icons, environmental and abstract sounds as scenarios became more urgent. Again, speech showed a flat function for urgency. Almost identical patterns for urgency and pleasantness were seen for the second weekly session, and so these are not shown graphically. A 2-way ANOVA with factors of score type (pleasantness, urgency) and sound group (speech, auditory icons, environmental, abstract) for week 1 showed no significant effect of scoring type or sound group (ps > .1), but a significant interaction (F(3, 36) = 7.503, p < .01). Similarly, in week 2 there was no significant effect of score type or sound group (Ps > .1), but there was a significant interaction (F3, 36) =

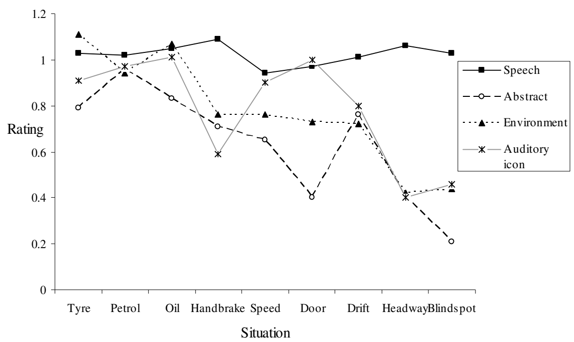

Figure 3. *Mean free-modulus magnitude scores for pleasantness in week 1.*

8.396, p < .01). Separate 2-way ANOVAs showed no difference across weeks 1 and 2 for either urgency or pleasantness scores (ps > .1).

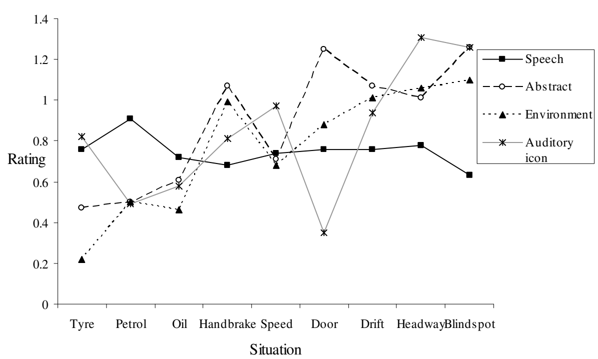

Figure 4. *Mean free-modulus magnitude scores for urgency in week 1.*

4. DISCUSSION

Sound types in this experiment were mapped onto particular driving scenarios and participant speed of responding and accuracy were measured. Speech and auditory icons produced both the faster response times and greatest accuracy. Abstract sounds produced the slowest response times and least accuracy. Environmental sounds showed an intermediate pattern of performance for accuracy but the response times were similar to the abstract sounds. In addition, a free modulus magnitude estimation measure was used to assess both perceived urgency of the sounds and speech, and their pleasantness. The results are easily summarised. Speech alone showed an insensitivity to both measures: utterances were similarly and consistently rated as pleasant, but also of intermediate urgency (that is, speech sounds did not differ according to situational urgency). On the other hand the three other sound types both mapped successfully onto the appropriate situational urgency and showed a consistent relationship: sounds mapped onto highly urgent scenarios were also judged less pleasant. It is perhaps not surprising that the environmental and abstract sounds were judged more urgent with increasing situational urgency: they were after all chosen according to how appropriate they were in according such urgency; but the results for the auditory icons offer a suggestion that their urgency reflected inherent learned or 'direct' [4] urgency (as opposed to reflecting only their acoustic characteristics).

Others [17] have suggested that alarms should evoke a fundamental link to their referent in terms of Gibsonian ‘affordances’ because, although the acoustic properties of a signal are undoubtedly important, they can be overpowered by the direct cognitive mapping of sounds to the actions they signify. Of course, one challenge for researchers is to attempt to identify those acoustic features of sounds which 'carry' easily learned mappings. The effectively ceiling performance for accuracy and response times for the auditory icons in the present study over the two weekly sessions suggests that such mappings may be easily acquired, and recalled well (no decline in week 2 performance). With respect to speech, the response times and recall performance is encouraging too, and contrasts with some former reports that indicate performance penalties for speech-based displays over time in a control-room task simulation [27]. The short three- or four-word utterances, beginning where possible with a word indicating the referent, is an important design consideration in the present study. The moderately high user acceptance for the speech displays (they were rated as more pleasant than most of the other sound displays here) is encouraging too. In summary, the results here point to the utility of auditory icons and speech as displays and alarms within vehicles.

As a concluding comment, it is undoubtedly the case that the very features that make abstract alarms attention-getting and probably annoying are welcome ones in emergencies which occur infrequently. So abstract sounds will probably continue to be employed as within-vehicle alarms, at least for those aspects of the interface where drivers' attention and action is paramount. Examples include low oil pressure and a punctured tyre. One promise of auditory icons is that they may serve as more frequently employed, and acceptable auditory displays which will not be confused with such infrequent, annoying alarms.

5. REFERENCES

[1] K. Catchpole, D. McKeown, and D. Withington, "Alerting, informing, and localizable auditory warnings," in *Engineering Psychology and Cognitive Ergonomics*, D. Harris, Ed., 1999, Vol. 4, Ashgate Aldershot, Brookfield, USA.

[2] B.A. Colavita, "Human sensory dominance," *Perception and Psychophysics*, vol. 16, pp. 409 – 412, 1974.

[3] B.H. Deatherage, "Auditory and other sensory forms of information processing," in *Human Engineering Guide to Equipment Design*, H.P. Van Cott and R.G. Kincade, Eds., 1972, American Institutes for Research: Washington D.C.

[4] P. Keller and C. Stevens, "Meaning from environmental sounds: Types of signal-referent relations and their effect on recognizing auditory icons, *Journal of Experimental Psychology: Applied*, vol. 10, pp. 3-12, 2004.

[5] R. Graham, "Use of auditory icons as emergency warnings: evaluation within a vehicle collision avoidance application," *Ergonomics*, vol. 42, pp. 1233-1248, 1999.

[6] D. Begault, *3-D sound for virtual reality and multimedia*. Academic Press, Cambridge, M.A., 1994.

[7] N.A. Stanton and J. Edworthy, "Auditory warning affordances," in *Human Factors in Auditory Warnings*, N.A. Stanton and J. Edworthy, Eds., 1999, Ashgate, Aldershot, Brookfield, USA, pp. 113-127.

[8] J.J. Gibson, J.J., *The Ecological Approach to Visual Perception*. New York, Wiley, 1979.

[9] S.D. Jones, and S.M. Furner, "The construction of audio icons and information cues for human-computer dialogues," in *Contemporary Ergonomics*, E. D. Megaw, Ed., 1989, Taylor and Francis, London.

[10] P.A. Lucas, "An evaluation of the communicative ability of auditory icons and earcons," in Proceedings 2nd International Conference on Auditory Display, G. Kramer, Ed., 1995, pp. 121 – 128.

[11] E.J. Hellier, J. Edworthy, and I. Dennis, "A comparison of different techniques for scaling perceived urgency," *Ergonomics*, vol. 38, pp. 659 – 670, 1995.

[12] E.C. Haas and J.G. Casali, "Perceived urgency and response time to multi-tone and frequency-modulated warning signals in broadband noise," *Ergonomics*, vol. 38, pp. 2313 – 2326, 1995.

[13] J.L. Burt, D.S. Bartolome, and D.W. Burdette, "A psychophysiological evaluation of the perceived urgency of auditory warning signals," *Ergonomics*, vol. 38, pp. 2327 – 2340, 1995.

[14] K.L. Momtahan, "Mapping of psychoacoustic parameters to the perceived urgency of auditory warning signals," Unpublished Masters thesis, University of Carleton, Ottawa, Ontario, Canada, 1990.

[15] J. Edworthy, S. Loxley, and I. Dennis, "Improving auditory warning design: relationship between warning sound parameters and perceived urgency," *Human Factors*, vol. 33, pp. 205 – 231, 1991.

[16] E.J. Hellier, J. Edworthy, and I. Dennis, "Improving auditory warning design: quantifying and predicting the effects of different warning parameters on perceived urgency," *Human Factors*, vol. 35, pp. 693 – 706, 1993.

[17] A. Guillaume, L. Pellieux, V. Chastres, and C. Drake, "Judging the urgency of nonvocal auditory warning signals: Perceptual and cognitive processes," *Journal of Experimental Psychology: Applied*, vol. 9, pp. 196-212, 2003.

[18] R.D. Patterson, *Guidelines for auditory warning systems on civil aircraft*. Civil Aviation Authority, Paper 82017, 1982.

[19] H. Lazarus and H. Höge, "Industrial safety: Acoustic signals for danger situations in factories," *Applied Ergonomics*, vol. 17, pp. 41 – 46, 1986.

[20] D.A. Laird, and K. Coye, " Psychological measurements of annoyance as related to pitch and loudness," *Journal of the Acoustical Society of America*, vol. 1, pp. 158 – 163, 1929.

[21] W. Spieth, "Annoyance threshold judgements of bands of noise," *Journal of the Acoustical Society of America*, vol. 28, pp. 872 – 877, 1956.

[22] S. Dornic and T. Laaksonen, "Continuous noise, intermittent noise, and annoyance," *Perceptual and Motor Skills*, vol. 68, pp. 11 – 18, 1989.

[23] J.P. Bliss, "Investigations of alarm mistrust under conditions of varying alarm and ongoing task criticality," in *Human Factors in Auditory Warnings*, N.A. Stanton and J. Edworthy, Eds., 1999, Ashgate, Aldershot, Brookfield, USA, pp. 113-127.

[24] S. Hirst and C. Johnson, "Auditory displays for in-vehicle systems," CED3 Report Number 1, HUSAT, Loughborough University of Technology, November, 1992.

[25] T. Engen, "Scaling methods," in J. Kling and L. Riggs, Eds., *Experimental Psychology*, 1971, London: Methuen, pp. 47-86.

[26] S.S. Stevens, "The direct estimation of sensory magnitudes: loudness," *American Journal of Psychology*, vol. 69, 1-25, 1956.

[27] N.A. Stanton, C. Baber, C., "Speech-based alarm displays," in N, *Human factors in auditory warnings*, A. Stanton and J. Edworthy, Eds., 1999, Ashgate, Aldershot, Brookfield, USA, pp. 243-261.